PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 :		(11) International Publication Number: WO 87/ 03910
C12Q 1/02, 1/68, G01N 33/53 A61K 37/10, 39/395, C07H 15/12	A1	(43) International Publication Date: 2 July 1987 (02.07.87)

(21) International Application Number: PCT/US86/02798

(22) International Filing Date: 23 December 1986 (23.12.86)

(31) Priority Application Number: 812,955

(32) Priority Date: 24 December 1985 (24.12.85)

(33) Priority Country: US

(71) Applicant: MASSACHUSETTS INSTITUTE OF TECHNOLOGY [US/US]; 77 Massachusetts Avenue, Cambridge, MA 02139 (US).

(72) Inventors: PENMAN, Sheldon ; 123 Fuller Street, Brookline, MA 02146 (US). FEY, Edward, G. ; 80 Commercial Street, Boston, MA 02109 (US).

(74) Agent: PABST, Patrea, L.; Hamilton, Brook, Smith & Reynolds, Two Militia Drive, Lexington, MA 02173 (US).

(81) Designated States: AT (European patent), BE (European patent), CH (European patent), DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

Published
With international search report.

(54) Title: CELLULAR METHOD FOR DETERMINING TISSUE OF ORIGIN

(57) Abstract

A biochemical procedure for identification and characterization of cells in a biopsy or sample of a body fluid. The method can be used to determine cell type, i.e. epidermal, neuronal; tissue of origin, i.e. breast tissue, liver tissue; and degree of abnormality. The procedure can also be used to make antibodies and hybridization probes to detect cell or tissue specific antigen and nuclear matrix associated nucleic acids in cellular material and body fluids. The procedure is based on the isolation and analysis of the components of a specific subcellular protein fraction referred to here as the "nuclear matrix". The nuclear matrix includes proteins and nuclear matrix associated DNA specific to different cell types. These proteins and nucleic acids are altered or new ones expressed as a result of viral infection, genetic defects or malignancy. The method has a number of important clinical applications in determining tissue type, tissue of origin, degree of malignancy and extent of metastasis in cancer patients; in detecting and analyzing chromosomal deficiencies or genetic defects, especially in cells obtained by amniocentesis; in identifying viral or other infections; and in measuring the extent and location of cell damage, particularly in patients with localized cell damage or autoimmune disease. The isolated nuclear matrix proteins are also useful in screening for drugs binding to and affecting the nuclear matrix.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**FR**	France	**ML**	Mali
AU	Australia	**GA**	Gabon	**MR**	Mauritania
BB	Barbados	**GB**	United Kingdom	**MW**	Malawi
BE	Belgium	**HU**	Hungary	**NL**	Netherlands
BG	Bulgaria	**IT**	Italy	**NO**	Norway
BJ	Benin	**JP**	Japan	**RO**	Romania
BR	Brazil	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SN**	Senegal
CH	Switzerland	**LI**	Liechtenstein	**SU**	Soviet Union
CM	Cameroon	**LK**	Sri Lanka	**TD**	Chad
DE	Germany, Federal Republic of	**LU**	Luxembourg	**TG**	Togo
DK	Denmark	**MC**	Monaco	**US**	United States of America
FI	Finland	**MG**	Madagascar		

CELLULAR METHOD FOR DETERMINING TISSUE OF ORIGIN

Reference to Related Application

This application is a continuation-in-part of our copending U.S. patent application Serial No. 812,955, filed December 24, 1985, entitled "Method for Determining Tissue of Origin and Degree of Malignancy of Tumor Cells".

Background of the Invention

The U.S. Government has certain rights in this invention by virtue of National Institute of Health Grant Numbers 5R01 CA08416-20 and 1R01 CA37330-01 and National Science Foundation Grant Number PCM 8309334.

Diagnosis of viral infection, cancer, chromosomal defects or autoimmune disease is often difficult and inexact. Heretofore, determining the properties of tumor cells or other abnormal cells has been the province of the clinical pathologist. Diagnosis is generally based on the morphology of the cells in histological preparations. Such diagnosis has serious limitations and cannot always distinguish tumor type and tissue of origin. There is a great need for alternative means of identifying cell type and stage of malignancy or abnormality. Chromosomal defects can be detected only in the case of gross morphological defects or where the proteins encoded by the missing or defective genes are known and can be assayed for. Viral infections can usually be diagnosed only by measuring antibody levels, examination of cells, and presence of clinical symptoms.

There is also a need for a means of identifying the site of tissue damage, not only in autoimmune diseases where the target cell is known, but also, for example, in bladder infection

or myocardial infarction, where cell degradation products might be released into the urine or bloodstream.

Attempts have been made to determine cell type by analysis of the protein composition of whole cell extracts. However, these extracts contain a number of different proteins, of which the vast majority do not vary between cell types. Even with techniques providing increased resolution between proteins, such as the more recent methods of two dimensional gel electrophoresis, such efforts have largely failed to find meaningful differences in proteins that could reliably serve as a basis for cell and tissue type identification. Even where there might be a change in proteins due to infection or malignancy, existing methods fail to differentiate the new or altered proteins from background proteins.

All cells, both plant and animal, have a nucleus surrounded by the cell cytoplasm. The nucleus contains the cellular DNA complexed with protein and termed chromatin. The chromatin, with its associated proteins, constitutes the major portion of the nuclear mass. The chromatin is organized by the internal skeleton of the nucleus, referred to here as the nuclear matrix. Although it is known that the nuclear matrix contains cell-type specific proteins and a small percentage of the total DNA, current electron microscopy techniques do not image the matrix and there are no reliable methods of separating specific non-chromatin matrix proteins from the much larger quantity of chromatin proteins. The methods that have been used either separated out the soluble proteins and discarded the insoluble proteins, or used high salt (2 M NaCl or 2 M LiCl) and other harsh solvents to extract these nuclear matrix proteins.

There has been limited success in biochemically identifying cells by analyzing the protein composition of a

subcellular fraction consisting of the intermediate filaments. The intermediate filaments are proteins present in all cells and can be used to discriminate between five major classes of cells: epithelial, neuronal, glial, muscle and mesenchymal cells such as fibroblasts. Labeled antibodies to these proteins only serve to distinguish among these broad cell classes, although some further discrimination is possible with epithelial cells. In epithelial cells, the intermediate filaments, the cytokeratins, are complex and differ between many types of epithelia. However, the cytokeratins can only be specifically distinguished chemically: antibodies have so far proven unable to discriminate precisely among more than a few of the cytokeratins. There are also reports that the cytokeratins are altered in some malignancies making their use for fine discrimination uncertain.

It is therefore an object of the present invention to provide a method for determining tissues of origin and stages of abnormality of tumor cells, virally transformed cells, and genetically defective or deficient cells.

It is another object of the present invention to provide a method for specifically identifying the tissue of origin and extent of tissue injury in disease states involving cellular destruction, including autoimmune diseases, myocardial infarction, and bacterial infection.

It is a further object of the invention to provide a method which can be performed relatively easily and quickly.

It is a still further object of the invention to provide a method for analyzing tissue of origin which is highly reproducible, objective and can be standardized.

It is yet another object of the present invention to provide biochemical methods including antibodies and hybridization probes for isolation and analysis of cell or tissue specific nuclear matrix proteins and nuclear matrix associated DNA with much greater resolution than previous methods.

Summary of the Invention

A procedure for analyzing tissue of origin and state of abnormality of genetically defective or deficient cells, virally infected cells, or tumor cells with far greater discrimination than previous methods of identification. The procedure also provides a means for identifying or imaging the tissue of origin and degree of cell damage in disease states causing cellular destruction and release of degradation products into body fluids, including autoimmune diseases.

In one embodiment of the method of the present invention, the cell nucleus is isolated, the cytoskeleton proteins and chromatin removed, the nuclear matrix isolated, and the "interior" and "exterior" components of the nuclear matrix separated. The "nuclear matrix", a specific fraction of cell protein constituting less than five percent of the total protein and six percent of the total DNA of the cell, contains many proteins that differ according to cell type and is highly enriched with cell type-specific antigens including highly cell type- and transformation-specific proteins that cannot be detected using prior art procedures. The method makes use of the unique properties of the nuclear matrix to achieve complete separation from all other cell constituents. The method is simple, rapid, reproducible, achieves a high degree of matrix purity, is applicable to essentially all types of cells, does not disrupt matrix morphology, and yields most of the biochemically significant matrix components. The method for isolating the nuclear matrix proteins is summarized as follows:

1. Isolation and separation of cells.

2. Separation of soluble cell proteins from the nucleus and cytoskeleton by extraction of membrane lipids and soluble proteins with a non-ionic detergent-physiological salt solution.

3. Separation of cytoskeleton proteins from the nucleus by solubilization of the insoluble cell material from step 2 in either 0.25 M ammonium sulfate pH 6.8, a detergent-sodium deoxycholate solution, or other gentle extraction buffer.

4. Separation of chromatin from the nuclear matrix by digestion of the insoluble material from step 3 with DNAase I and RNAase in a physiological buffer and elution of the DNA-containing nucleosomes with 0.25 M ammonium sulfate solution buffered to pH 6.8 or other gentle extraction buffer.

5. Separation of the "interior" and "exterior" nuclear matrix proteins by dissolution of the insoluble material from step 4 in a buffer containing between 5 and 10 M urea, preferably 8 M urea, or other suitable solubilizing agent, and aggregation of the exterior proteins by dialysis into physiological buffer.

Further purification of the fractions selectively enriched in the "interior" and "exterior" proteins can be performed using HPLC, FPLC, chromatofocusing, and other methods known to those skilled in the art.

In a variation of this procedure, the cytoskeleton proteins and chromatin are removed together by digesting the insoluble material from step 2 with DNAase and RNAase, then extracting with 0.25 M ammonium sulfate at pH 6.8.

In another embodiment of the method, the nuclear matrix associated DNA is isolated and analyzed, alone or in conjunction with the nuclear matrix proteins. The method consists of treating the insoluble material from step 2 with DNAase then 0.25 M ammonium sulfate at physiological pH to remove the chromatin. A phenol extraction and/or centrifugation in $CsCl_2$ is then performed to remove any remaining protein. In a variation of this method, the insoluble material from step 2 is digested with a restriction enzyme in the appropriate buffer, the chromatin extracted with 0.25 M ammonium sulfate pH 6.8, then any remaining protein removed by phenol extraction and/or centrifugation in $CsCl_2$.

Because of the degree of matrix purification obtained with this procedure, several previously unknown properties of the nuclear matrix proteins have been discovered which are useful in the clinical diagnosis of tumor cells, virally infected cells, cells having chromosomal defects or genetic deficiencies, or with diseases involving specific cell damage such as autoimmune disease. The composition of the matrix proteins is different in every cell type and is further changed when the cell is transformed to the neoplastic state or has genetic aberrations. These proteins can be used to:

1. Identify general cell type (e.g., epithelial, neuronal, etc.).

2. Identify specific cell type (e.g., colon epithelium, breast epithelium, etc.). This

allows identification of the origin of metastasis or site of cell damage.

3. Supply pure antigens to be employed to produce tissue specific antibodies for use in immunoassays and in diagnosis and treatment of various diseases, wherein the antibodies can be conjugated with radioactive material, nuclear magnetic resonance contrast agents, boron derivatives, or biologically active agents.

4. Determine the nature and degree of malignancy, genetic deficiency or defect, or disease.

5. Screen drugs and other agents or compounds for binding to nuclear matrix proteins, including steroid receptors.

6. Identify DNA binding sites which may alter and regulate gene expression and cell differentiation.

Further characterization is achieved by analyzing and comparing the nuclear matrix associated DNA and, in particular, the restriction endonuclease fragments of nuclear matrix associated DNA. Hybridization probes can be prepared from the isolated, synthesized or cloned DNA which may be employed in a variety of assays for cell type, tissue of origin, malignancy, infection, or genetic abnormality.

Brief Description of the Drawings

Fig. 1 is a schematic of the isolation and analysis of nuclear matrix proteins and associated DNA according to the present invention;

Fig. 2a is a comparison of two dimensional electropherograms (pI vs. m.w.), both actual and diagrammatic, of nuclear matrix proteins from human lung (black circles) and adrenal cortex (white circles) where proteins common to both are shown as half black circles; and

Fig. 2b is a comparison of two dimensional electropherograms (pI vs. m.w.), both actual and diagrammatic of nuclear matrix proteins from human colon (black circles) and bladder (white circles), where proteins common to both are shown as half black circles.

Detailed Description of the Invention

The present invention is a method for isolating and identifying nuclear matrix proteins and nuclear matrix associated DNA unique to specific cells of a particular tissue type, some of which are altered by malignancy, viral or bacterial infection, or genetic defect. The sequential extraction yields subfractions of biochemically distinct cellular proteins and DNA as well as morphologically distinct nuclear matrix structures.

In one embodiment of the invention, the nuclear matrix is purified from a cell suspension prepared from a tissue biopsy or blood sample, separated into its "interior" and "exterior" fractions, and then analyzed by two-dimensional gel electrophoresis. Prior to the present invention, no one has been able to separate these fractions nor to extract the nuclear matrix proteins in a way which did not cause extensive degradation. The nuclear matrix proteins account for less than approximately five percent of the total cell protein. Many of the nuclear matrix proteins change when cells are transformed to malignancy. The "interior" proteins represent less than one percent of the total protein, with the "exterior" or

"intermediate filament" proteins making up the remainder. The designations "interior" and "exterior" are somewhat arbitrary but in general refer to the localization of the proteins within the nucleus. The interior matrix proteins reflect the structure and composition of the cell being examined. The exterior matrix proteins, by virtue of containing the intermediate filaments, reflect the class of the cell being examined (neuronal, epithelial, etc.).

Separating the exterior proteins from the interior matrix proteins has two important benefits. One, the sensitivity and specificity with which the interior matrix proteins can be identified is greatly enhanced since the exterior matrix proteins amount to greater than one-half of the total protein of the fraction and their presence can obscure less abundant but important interior matrix proteins. Secondly, while inadequate for the purpose of determining tissue type, the exterior proteins themselves provide two analytic tools. The intermediate filament proteins are useful for quick identification of the class of cell being analyzed and the intermediate filament-associated proteins can assist in the identification of disease.

Proteins isolated using the procedure of the present invention are useful as immunogens for the preparation of antibodies. When these antibodies are conjugated with colorimetric, immunological (such as labeled anti-antibody), fluorescent, enzymatic or radioactive labels, they can be used as cytological stains on histological sections or for analysis of body fluids and proteins separated by two dimensional gel electrophoresis to provide useful diagnostic information. The antibodies can detect the presence of tumor or viral antigens, abnormal proteins or the absence thereof due, for example, to a defective chromosome or genetic deficiency, and proteins released

during cell destruction, as in some types of infection, myocardial infarction, and autoimmune diseases. Antibodies labeled with radioactive material, nuclear magnetic resonance contrast agents, or boron or boron compounds are particularly useful for diagnostic imaging (detection or monitoring) or treatment of various diseases. Antibodies with biologically active agents, including drugs, biological modifiers, angiogenesis or anti-angiogenesis factors, can be used to deliver the agent to a selected location within an individual at a faster rate and higher efficacy of delivery than compounds delivered by presently known methods.

These nuclear matrix proteins are also useful in directly comparing the efficacy of a number of compounds as candidates for treatment of a specific disease (drug screening). For example, steroids appear to bind to the nuclear matrix via some nuclear matrix receptor mechanism. Binding of estrogen or phenol red can stimulate cell proliferation. Binding to the nuclear matrix receptors or proteins seems to modulate specific gene expression. Drugs which interact with and affect the nuclear matrix may have a number of therapeutic applications. The present invention provides a means for comparing the effects of these drugs. The nuclear matrix proteins are isolated, as previously described, exposed to the compound to be tested, and the presence of binding measured. Binding assays are well known to those skilled in the art. Once a compound is found to bind in vitro, it may be tested in vivo for a specific effect on gene expression, proliferation or inhibition. However, the advantage of a rapid, in vitro assay is obvious--only those drugs which bind need to undergo further testing.

The nuclear matrix associated DNA extracted by this procedure has similar uses. The procedures for isolating the

interior and exterior nuclear matrix proteins and nuclear matrix associated DNA are as follows:

Cell Preparation

The procedure employs single cells in a suspension. Cells in blood samples or from cell culture are already separate. Tissues obtained by biopsy are dispersed by mild mechanical homogenization followed by digestion with a proteolytic enzyme such as collagenase or trypsin. These enzymes digest connective fibers between cells without affecting the interior contents. In some cases it is desirable to perform an initial, partial separation of cell types. A rapid cell separation is effected by centrifugation, an inert density gradient, or by other means for rapidly separating cells without chemical interaction.

Purification of the Nuclear Matrix Proteins and Associated DNA

The nuclear matrix is separated from other cell constituents by a series of sequential extractions. The cell suspension is exposed to the extraction solution for one to two minutes and then the insoluble material separated by centrifugation (approximately one to two minutes at 1000 g), filtration (pore size approximately 5 microns), or other method known to those skilled in the art. The sequential fractionation is shown schematically in Fig. 1. One key feature of this extraction process is that harsh extraction solvents such as 2 M NaCl are avoided. A second is the use of urea or other suitable solubilizing agent to dissolve the interior and exterior proteins, followed by repolymerization of the exterior proteins. The steps are as follows:

1. Removal of soluble cell proteins.

The soluble proteins 14, amounting to 70% of the cell

mass, are removed by extracting the intact cell 10 with a non-ionic detergent solution 12, such as 0.5% Triton X-100. The non-ionic detergent in a buffer at physiological pH and ionic strength extracts first the membrane lipids and then the soluble proteins. The detergent solubilizes lipids without denaturing proteins and thus avoids disturbing the integrity of the cell structures. The physiological salt solution is essential to maintain the morphology of the cytostructure and to prevent the removal of structural elements. An example of a useful buffer is 100 mM NaCl, 300 mM sucrose, 10 mM PIPES [pH 6.8], 3 mM $MgCl_2$, 0.5% Triton X-100, 1.2 phenylmethylsufonylfluoride at 0°C.

Two dimensional gel analysis of the soluble fraction reveals a complex, dense pattern of major proteins. While many proteins appear to be unique in this fraction, the density of protein spots on the gel precludes more precise analysis.

The remaining skeletal framework, usually masked in conventional Epon-embedded thin section, can be visualized in three dimensions as an unembedded whole mount using the method of Fey et al, in the J. Cell Biol., 98,1973-1984 (1984). For scanning electron microscopy, cells are grown on glass coverslips and fractionated using the method of the present invention. Cells are fixed at various stages of fractionation in the appropriate buffer containing 2.5% gluteraldehyde at 0°C for 30 minutes followed by rinsing in 0.1 M sodium cacodylate and then 1% OsO_4 in 0.1 M Na cacodylate for 5 minutes at 0°C. The cells, still on coverslips, are dehydrated through an ethanol series, dried through the CO_2 critical point and sputter-coated with gold-palladium. The samples are examined in the lower stage of a scanning electron microscope. Transmission electron microscopy is done on cells grown on gold grids which were previously covered with formvar and coated with carbon. Cells are fixed in 2.5% gluteraldehyde and processed as above.

2. Removal of cytoskeleton proteins.

The cytoskeleton proteins 20, consisting of the dense cytoplasmic filament networks, intercellular junctional complexes and apical microcellular structures, are next separated from the nucleus by selective solubilization of the cytoskeleton proteins amounting to 20% of cell protein mass, with either a 0.25 M ammonium sulfate solution 16 buffered to pH 6.8, for example, 0.25 M ammonium sulfate, 0.3 M sucrose, 10 mM PIPES [pH 6.8], phenylmethylsulfonyl fluoride, 0.5% Triton X-100, or with a 1% Tween-40, 0.5% sodium deoxycholate solution 18. The nucleus with all of the intermediate filaments still attached, amounting to about 5% of the total cellular protein, remains.

3. Removal of the chromatin proteins.

The chromatin proteins 26, whose association with the nucleus depends on the integrity of DNA and RNA, is next separated from the nuclear matrix. The nucleus is first digested with DNAase and RNAase in near physiological digestion buffer 22, for example, 50 mM NaCl, 0.3 M sucrose, 10 mM PIPES [pH 6.8], 3 mM $MgCl_2$, 0.5% Triton X-100, 1.2 mM phenylmethylsulfonyl fluoride with 100 micrograms bovine pancreatic DNAse (EC 3.1.4.5, Worthington Biochemical Corp., Freehold, NJ) and 100 micrograms/ml pancreatic RNAse A (EC 3.1.4.22., Sigma Chemical Co., St. Louis, MO). The enzyme cuts DNA principally between the basic units of chromatin or nucleosomes. At this stage, the DNA remains completely in the nucleus in the form of individual nucleosomes. The DNA-containing nucleosomes are then eluted for approximately 5 minutes at 20°C using 0.25 M ammonium sulfate buffered to pH 6.8 24. Although 0.25 M ammonium sulfate is preferred, comparable buffers could be used.

4. Separation of the Interior and Exterior Proteins of the Matrix

The nuclear matrix consists of less than 5% of the cell protein mass.

The nuclear matrix is further divided into two distinct parts, termed the interior and exterior (or intermediate filaments). The exterior part consists of the intermediate filaments and intermediate filament associated proteins. Although the filaments are exterior to the matrix, in the cytoplasm, they are physically connected to the nuclear surface and behave physiologically as part of the nuclear matrix. They and their associated proteins amount to greater than one-half of the proteins in the matrix preparation.

Fractions selectively enriched in the interior and exterior matrix proteins are produced by completely dissolving the matrix proteins in a buffered 5 to 10 M urea solution 28, preferably 8 M, or other suitable solubilizing agent such as, for example, lithium diiodosalicylate, as required to completely dissolve the nuclear matrix proteins, and then dialyzing the proteins back into physiological buffer 30. The interior proteins 32 remain in solution. The intermediate filament proteins and associated proteins reassemble into large insoluble filaments 34 during dialysis to remove the solubilizing agent. The urea or comparable solubilizing agent is essential for dissolving the proteins, which are generally quite insoluble. Removal of the solubilizing agent, as by dialysis, is essential to reform and separate out the intermediate filament fractions.

In a variation of the method for isolating the nuclear matrix proteins described above, the cytoskeleton proteins 20 and chromatin 26 are removed together. The soluble proteins 14 are first removed by extraction with a non-ionic detergent buffered solution 12. The insoluble material is digested with DNAase and

RNAase in a buffered solution 22 then the cytoskeleton proteins 20 and chromatin 26 extracted with 0.25 M ammonium sulfate at physiological pH 28.

The selectively enriched fractions can be further separated by electrophoresis, "chromatofocusing" (BioRad Laboratories, Richmond, CA), HPLC and related technologies such as FPLC (fast protein liquid chromatography), isoelectrofocusing, and conventional ion exchange or affinity chromatography. Once specific proteins are separated, the proteins can be used to make monoclonal antibodies which can be bound to matrices for use in purification of additional protein, as well as in analysis.

Protein Analysis

The highly purified matrix proteins are analyzed using conventional two-dimensional acrylamide gel electrophoresis. The proteins are first separated in a pH gradient gel according to electrophoretic mobility or isoelectric point. This gel is then placed on a standard 10% acrylamide slab gel and the proteins separated according to molecular weight. One method of two dimensional gel electrophoresis is taught by P.H. O'Farrell in J. Biol. Chem. 250, 4007-4021 (1975) using an ampholyte gradient consisting of 90% pH 5-7 (0.4% ampholyte) and 10% pH 3-10 (1.6% ampholyte). The proteins form a pattern of spots, made detectable by silver staining or by autoradiography, which is diagnostic of the cell type and state of transformation or abnormalities. Equivalent ^{35}S cpm can be used to facilitate qualitative comparisons.

Immunoblot electrophoresis can also be used to identify proteins in the nuclear matrix - intermediate filament scaffold fraction, including vimentin, cytokeratins desmosomal proteins, and specific nuclear matrix proteins. The procedure for

immunoblot electrophoresis is as follows. One-dimensional polyacrylamide gels are run according to the method of Laemmli in Nature (Lond.), 227,680-685 (1970). Equal protein concentrations are loaded to compare individual fractions. The reaction of antibodies to protein bands are visualized on nitrocellulose paper according to the procedure of H. Towbin et al., Proc. Natl. Acad. Sci. USA, 76,4350-4354 (1979). Nitrocellulose strips are incubated for 12 hours in 2% hemoglobin in PBS, rinsed three times in PBS, and incubated for 2 hours at 20°C with antibody to the protein to be detected at the appropriate concentration. Excess antibody is removed by washing with PBS (four 20 minute washes). The strips are then incubated with goat anti-rabbit (or anti-mouse) IgG conjugated to horseradish peroxidase, washed in PBS (four times for a total of 80 minutes) and then developed in 0.4 mg/ml 4-chloro-1-napthol in 0.01% (vol/vol) H_2O_2 using the technique of R. Hawkes et al., in Anal. Biochem., 119,142-147 (1982).

The matrix preparation is biochemically and morphologically pure by several biochemical criteria. It retains most matrix specific constituents. Freedom from contamination permits a clear and detailed analysis of the matrix proteins by gel electrophoresis.

As shown in Fig. 1, nuclear matrix associated DNA 44 is isolated from the intact cell 10 by extraction of the soluble proteins 14 with a non-ionic detergent in a physiological buffered solution 12, solubilization of the chromatin 26 and cytoskeleton proteins 20 by digestion with DNAase in a physiological buffer 36 followed by extraction into 0.25 M ammonium sulfate at physiological pH 24, and removal of any remaining protein 42 by phenol extraction 38, centrifugation in a cesium chloride gradient 40 or other method known to those

skilled in the art. Digestion of the insoluble material with DNAase 36 instead of DNAase in combination with RNAase 22 results in the removal of approximately 94% of the cell DNA instead of approximately 98%.

In another embodiment of this procedure, also shown in Fig. 1, fragments of nuclear matrix associated DNA 48 are prepared by digestion of the insoluble cell material remaining after extraction of the soluble proteins 14 with one or more restriction endonucleases in the appropriate buffer 46, followed by extraction of the cytoskeleton proteins 20 and chromatin 26 in 0.25 M ammonium sulfate at physiological pH 24 and removal of any remaining protein 42 by phenol extraction 38, cesium chloride centrifugation 40, or other method known to those skilled in the art.

The two dimensional gel electropheresis and other purification techniques yield pure proteins for use in the production of antibodies. Polyclonal or monoclonal antibodies may be preferred, depending on the antigen to be detected and the technique utilized. Monoclonal antibodies are secreted by hybridoma cell lines produced by established immunization and fusion protocols known to those skilled in the art, such as the method of G. Galfre et al., in Nature (Lond.) 266,550-552 (1977).

Cell type specificity of matrix proteins has been demonstrated using laboratory cultured cell lines. These include common laboratory lines such as primary and established fibroblasts, HeLa cells, etc. Most relevant to clinical applications are the results from a number of carcinoma lines derived from human biopsy material. As shown in Fig. 2, these include human colon, lung, adrenal cortex and bladder cell lines. Although the electropherogram patterns are markedly different in

different cell types, each cell type containing unique as well as common proteins, the pattern for each cell type is specific and reproducible. Similar results can be obtained with animal tissue, having been demonstrated using mouse tissue.

In addition to identifying the tissue of origin of the nuclear matrix proteins, the matrix protein patterns reflect cell transformation, viral infection, and genetic defects. This has been demonstrated in spontaneous transformation of rodent primary fibroblasts, viral transformation of established rodent fibroblasts, transformation of a kidney cell line by transfection with the ras oncogene, and transformation of a kidney cell line by the ultimate carcinogen, BAP diole epoxide, using a number of cloned cell lines in each case. Each showed different and marked variation in their matrix protein pattern, but retained sufficient information to determine the cell of origin.

The type and degree of transformation or infection are also closely correlated with changes in matrix protein composition. Of particular interest are the qualitative differences between transformation by the complete carcinogen and by ras gene transfection. Six cloned isolates of cells transformed by ras transfection and 10 isolates of the carcinogen-transformed cells were analyzed. A brief summary of the results is:

a. Transformation by chemical carcinogen leads to the appearance in the matrix of 12 to 15 new or previously undetected proteins.

b. Transfection by the ras gene results in the loss of about 6 proteins and, in some cases, the appearance of 2 or 3 new proteins.

There appears to be a correlation between the degree of morphological aberrance and the number of matrix protein changes.

The data from transformed cells reveals different types of transformation events with characteristic signatures.

Analysis of the nuclear matrix associated DNA provides further information for determining the cell type, tissue of origin, and degree of abnormality of cells. "Actively" transcribed DNA in a cell is usually associated with the nuclear matrix proteins. This DNA represents approximately 6% of the total cell DNA. Approximately one-third (2% of total DNA) is directly bound to the protein portion of the matrix and approximately two-thirds (4% of total DNA) is bound to the RNA component of the nuclear matrix. Different quantities of DNA as well as specific sequences of DNA will be associated with the nuclear matrix and nuclear matrix proteins, depending on cell type and whether the cell is abnormal and to what degree the cell is abnormal. The isolated DNA can be analyzed and identified using gel electrophoresis and blot hybridization with probes specific for a particular unique sequence or repetitive sequence. Probes may be made by inserting the sequence of interest, either a synthetic sequence or a portion of the gene of interest, into a recombinant plasmid using methods known to those skilled in the art. The probes can be used to screen commercially available libraries, such as the GT11 library, for cDNA encoding nuclear matrix proteins. The library could also be screened using antibodies to the proteins. The probes can also be conjugated with a radiolabel, biotinylated, or crosslinked with psoralin and derivatives thereof for use as a double stranded probe and used in assays, imaging, isolation and identification procedures.

The isolated DNA can be inserted into a vector for expression of nuclear matrix proteins or expressed directly in a system such as one of the frog oocytes direct expression systems.

Analysis of either the proteins or the DNA provides a means to determine the presence of chromosomal defects or genetic deficiencies which might otherwise be undetectable. One application is in the analysis of cells obtained by amniocentesis. Another is in the identification and assessment of autoimmune diseases. Both the antibodies and hybridization probes can be used to analyze cellular materials and body fluids, both in vivo (tissue imaging) and in vitro, for nuclear matrix proteins. However, steps to concentrate or enhance the antigen levels in the body fluids may be required to insure adequate levels for detection. Once probes or antibodies are developed, for example, to carcinoma specific proteins, they can be labeled and used to rapidly screen either histological sections such as a pap smear, or the body fluid.

The extremely rapid, simple extraction and analytical procedure of the present invention provides a means for an objective determination of the tissue of origin of normal and abnormal cells and their degradation products, thereby indicating whether transformation or infection of normal cells has occurred, whether there has been metastasis and to what degree, and whether the individual has any autoimmune disease.

Although this invention has been described with reference to specific embodiments, variations and modifications of the method for isolating and diagnosing nuclear matrix proteins and associated DNA from cells of unknown tissue type or state of malignancy, infection or abnormality will be obvious to those skilled in the art. It is intended that such modifications and variations fall within the scope of the appended claims.

CLAIMS

1. A method for identifying and characterizing cellular materials and body fluids comprising:

a) isolating the interior and exterior nuclear matrix proteins from at least one type of cell of known origin;

b) separating and identifying isolated nuclear matrix proteins; and

c) comparing the nuclear matrix proteins from cells of unknown origin with the separated, identified nuclear matrix proteins of known origin.

2. The method of claim 1 further comprising

a) isolating the interior and exterior nuclear matrix proteins from cells of unknown origin;

b) separating and identifying isolated nuclear matrix proteins from cells of unknown origin; and

c) comparing the identified, separated nuclear matrix proteins from cells of unknown origin with the separated, identified nuclear matrix proteins of known origin.

3. The method of claim 1 wherein the nuclear matrix proteins are isolated by:

a) separating the soluble proteins from the nucleus and cytoskeleton by extracting eucaryotic cells with a non-ionic detergent solution at physiological pH and ionic strength;

b) separating the nucleus from the cytoskeleton proteins by solubilizing the cytoskeleton proteins in the insoluble material from step a;

c) separating the chromatin proteins from the nuclear matrix by digesting the insoluble material from step b with DNAse and RNAase and eluting the chromatin proteins with a buffered ammonium sulfate solution;

d) separating the interior and exterior proteins of the nuclear matrix by dissolving the insoluble material from step

c in a solubilizing agent and dialyzing the dissolved proteins back into a physiological buffer.

4. The method of claim 1 wherein the interior and exterior proteins of the nuclear matrix are solubilized in urea.

5. The method of claim 1 further comprising purifying the interior nuclear matrix proteins.

6. The method of claim 1 wherein the cytoskeleton proteins and chromatin proteins are removed together by digesting the insoluble material from step a with DNAase and RNAase.

7. The method of claim 1 further comprising isolating the nuclear matrix associated nucleic acids from at least one cell of known origin.

8. The method of claim 7 further comprising making antibodies or hybridization probes to the isolated, separated nuclear matrix proteins and associated nucleic acids.

9. The method of claim 8 further comprising labeling the anti-nuclear matrix protein antibodies or associated nuclear matrix nucleic acid hybridization probes and comparing the nuclear matrix proteins and associated nucleic acids from cells of unknown origin by reacting said proteins and nucleic acids with the labeled antibodies and hybridization probes to nuclear matrix proteins and nucleic acids of known origin.

10. The method of claim 8 further comprising selecting antibodies to nuclear matrix proteins, labeling said selected antibodies and injecting said labeled antibodies into an individual.

11. The method of claim 7 further comprising digesting the nuclear matrix associated nucleic acids with restriction enzymes.

12. The method of claim 7 wherein the nuclear matrix associated nucleic acid is isolated by:

a) separating the soluble proteins from the nucleus and cytoskeleton by extracting eucaryotic cells with a non-ionic detergent solution at physiological pH and ionic strength;

b) digesting the insoluble material from step a with DNAase and then eluting with a buffered ammonium sulfate solution; and

c) removing any remaining protein from the insoluble material in step b.

13. The method of claim 1 further comprising selecting the cells of known origin of step a from the group of cells consisting of normal cells, cells containing nucleotide sequences of viral origin or proteins derived thereof, malignant cells, and cells having chromosomal deficiencies or defects for the production of specific proteins.

14. A method for detecting nuclear matrix proteins and associated nucleic acids in samples of cellular material and body fluids comprising reacting labeled antibodies to isolated interior and exterior nuclear matrix proteins of known origin with the sample to be analyzed.

15. A method for detecting nuclear matrix associated nucleic acids in samples of cellular material and body fluids comprising reacting labeled hybridization probes to nuclear matrix associated nucleic acids of known origin with the sample to be analyzed.

16. A method for screening for compounds which bind nuclear matrix proteins or nucleic acids comprising

reacting selected compounds with isolated interior and exterior nuclear matrix proteins or nuclear matrix associated nucleic acids and determining the extent of binding.

17. Nuclear matrix proteins isolated from eucaryotic cells origin by:

a) separating the soluble proteins from the nucleus and cytoskeleton by extracting eucaryotic cells with a non-ionic detergent-salt solution at physiological pH and ionic strength;

b) separating the nucleus from the cytoskeleton by solubilizing the cytoskeleton proteins in the insoluble material from step a;

c) separating the chromatin proteins from the nuclear matrix by digesting the insoluble material from step b with DNAse and RNAase and eluting with a buffered ammonium sulfate solution; and

d) separating the interior and exterior proteins of the nuclear matrix by dissolving the insoluble material from step c in a solubilizing agent and dialyzing the dissolved proteins back into a physiological buffer.

18. Antibodies to the nuclear matrix proteins of claim 17.

19. cDNA and derivatives thereof encoding the isolated nuclear matrix proteins of claim 17.

20. Nuclear matrix associated nucleic acids isolated from eucaryotic cells origin by

a) separating the soluble proteins from the nucleus and cytoskeleton by extracting eucaryotic cells with a non-ionic

detergent solution at physiological pH and ionic strength;

b) digesting the insoluble material from step a with DNAase and then eluting with a buffered ammonium sulfate solution; and

c) removing any remaining protein from the insoluble material from step b.

21. Hybridization probes derived from the nuclear matrix associated nucleic acids of claim 20.

22. Restriction fragments of nuclear matrix associated nucleic acid prepared by

a) separating the soluble proteins from the nucleus and cytoskeleton by extracting eucaryotic cells with a non-ionic detergent solution at physiological pH and ionic strength;

b) digesting the insoluble material from step a with one or more restriction enzymes in the appropriate buffer and then eluting with a buffered ammonium sulfate solution; and

c) removing any remaining protein from the insoluble material in step b.

CELLULAR METHOD FOR DETERMINING TISSUE OF ORIGIN

Abstract of the Invention

A biochemical procedure for identification and characterization of cells in a biopsy or sample of a body fluid. The method can be used to determine cell type, i.e. epidermal, neuronal; tissue of origin, i.e. breast tissue, liver tissue; and degree of abnormality. The procedure can also be used to make antibodies and hybridization probes to detect cell or tissue specific antigens and nuclear matrix associated nucleic acids in cellular material and body fluids.

The procedure is based on the isolation and analysis of the components of a specific subcellular protein fraction referred to here as the "nuclear matrix". The nuclear matrix includes proteins and nuclear matrix associated DNA specific to different cell types. These proteins and nucleic acids are altered or new ones expressed as a result of viral infection, genetic defects or malignancy.

The method has a number of important clinical applications in determining tissue type, tissue of origin, degree of malignancy and extent of metastasis in cancer patients; in detecting and analyzing chromosomal deficiencies or genetic defects, especially in cells obtained by amniocentesis; in identifying viral or other infections; and in measuring the extent and location of cell damage, particularly in patients with localized cell damage or autoimmune disease. The isolated nuclear matrix proteins are also useful in screening for drugs binding to and affecting the nuclear matrix.

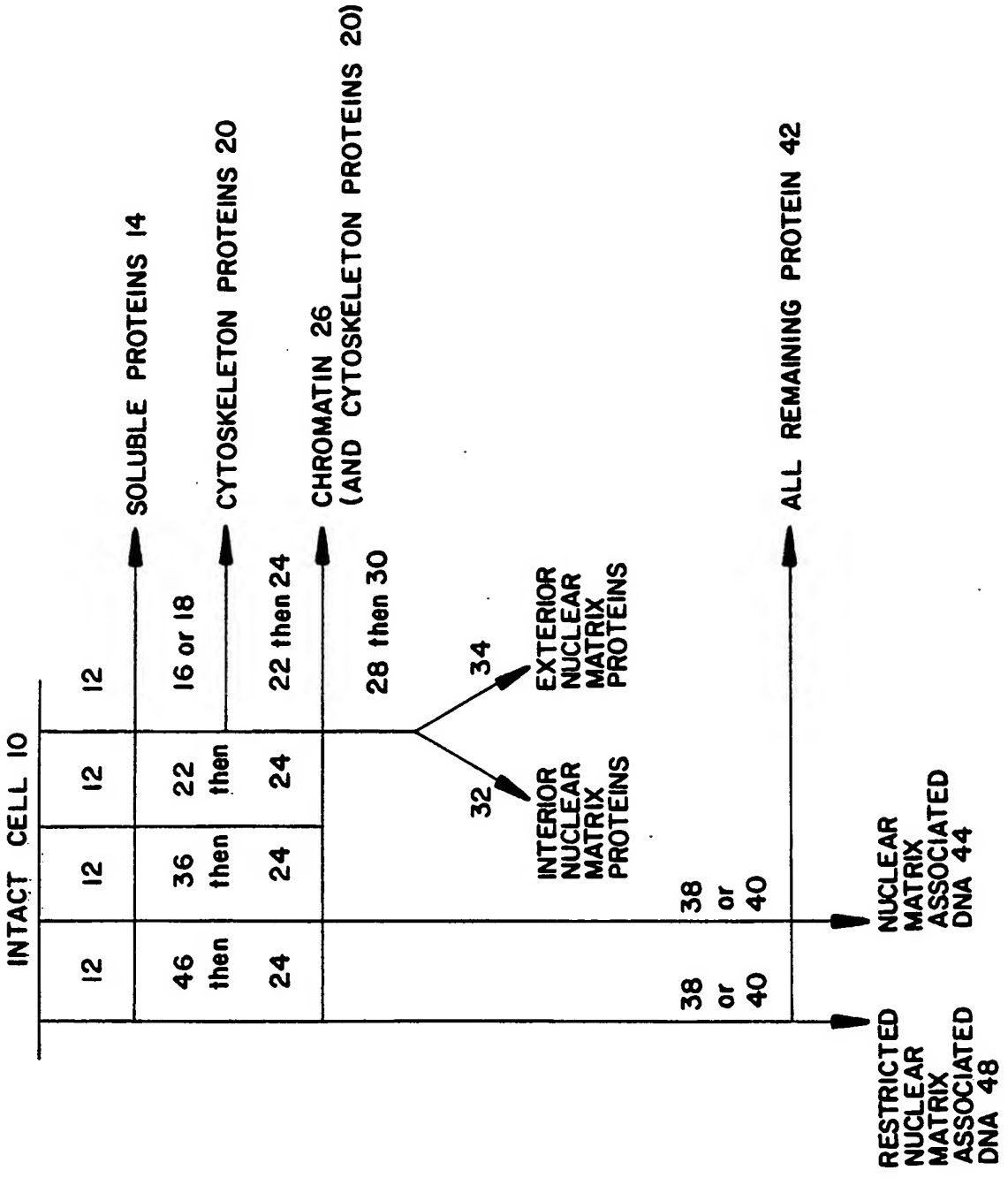
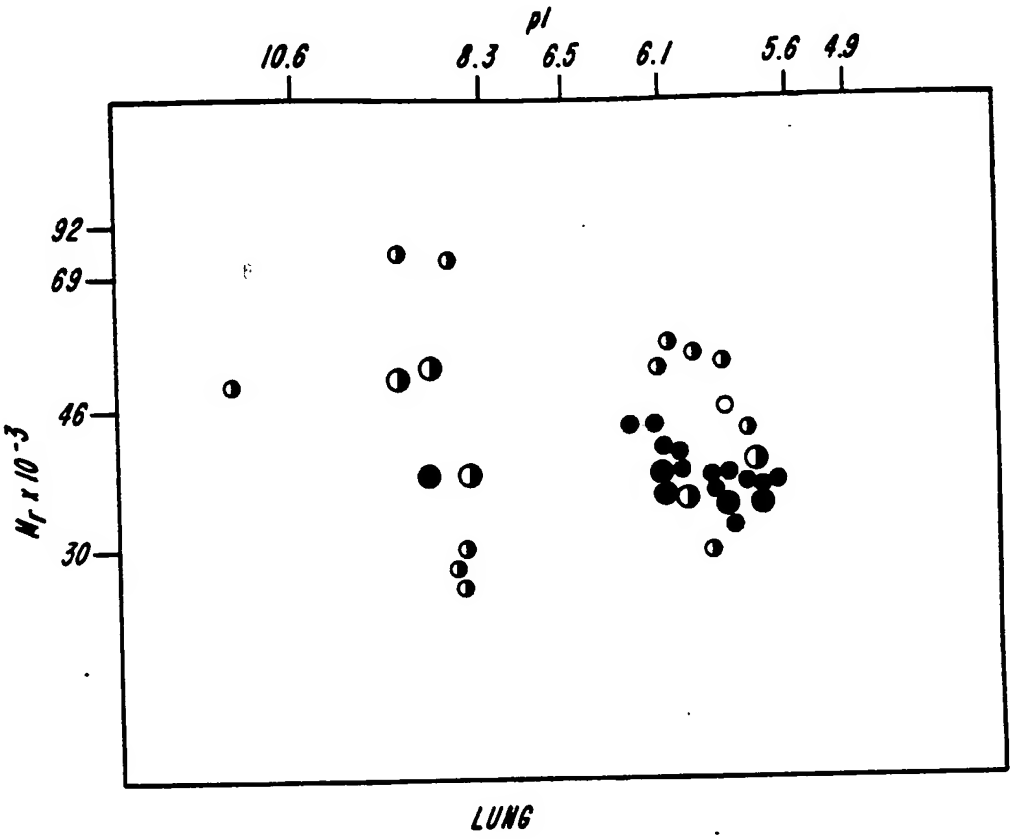

INTACT CELL 10
12
12
12
46 then 24
SOLUBLE PROTEINS 14
16 or 18
CYTOSKELETON PROTEINS 20
22 then 24
36 then 24
22 then 24
CHROMATIN 26 (AND CYTOSKELETON PROTEINS 20)
28 then 30
34
EXTERIOR NUCLEAR MATRIX PROTEINS
32
INTERIOR NUCLEAR MATRIX PROTEINS
38 or 40
38 or 40
ALL REMAINING PROTEIN 42
NUCLEAR MATRIX ASSOCIATED DNA 44
RESTRICTED NUCLEAR MATRIX ASSOCIATED DNA 48

FIG. 1



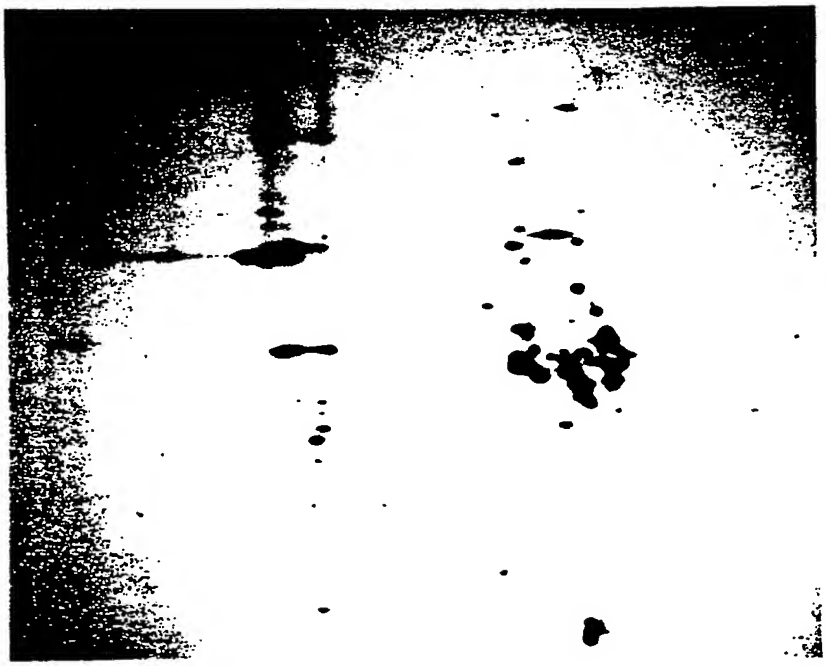


FIG. 2A

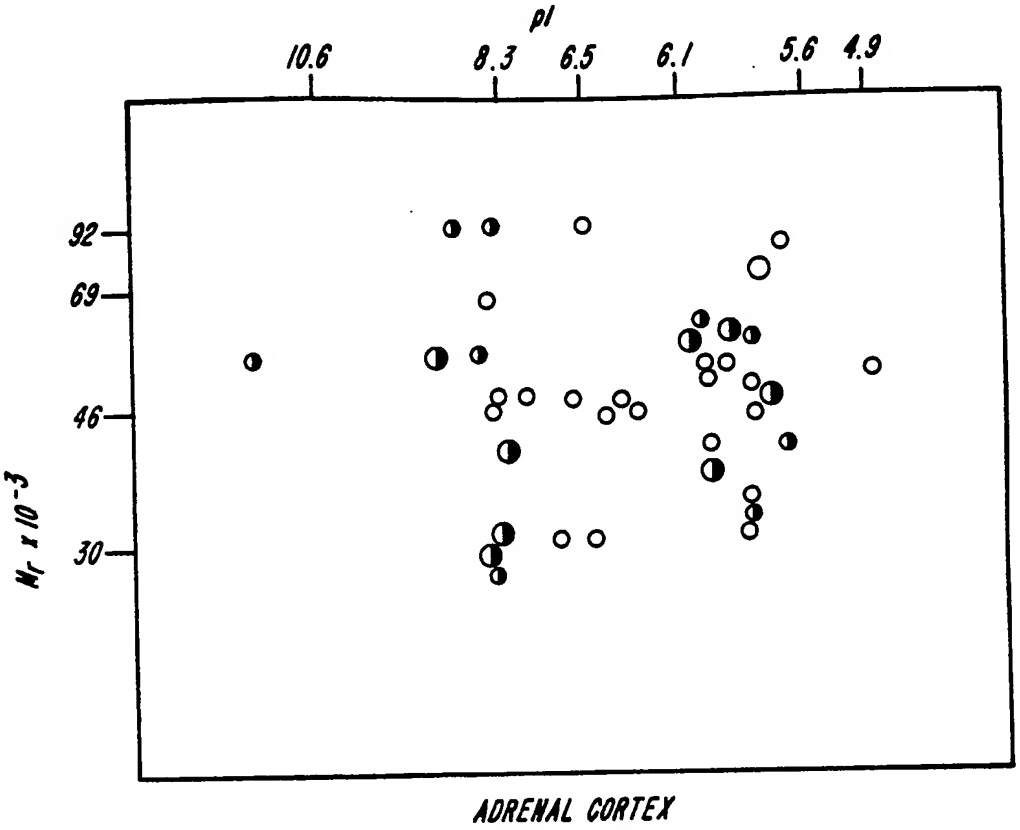
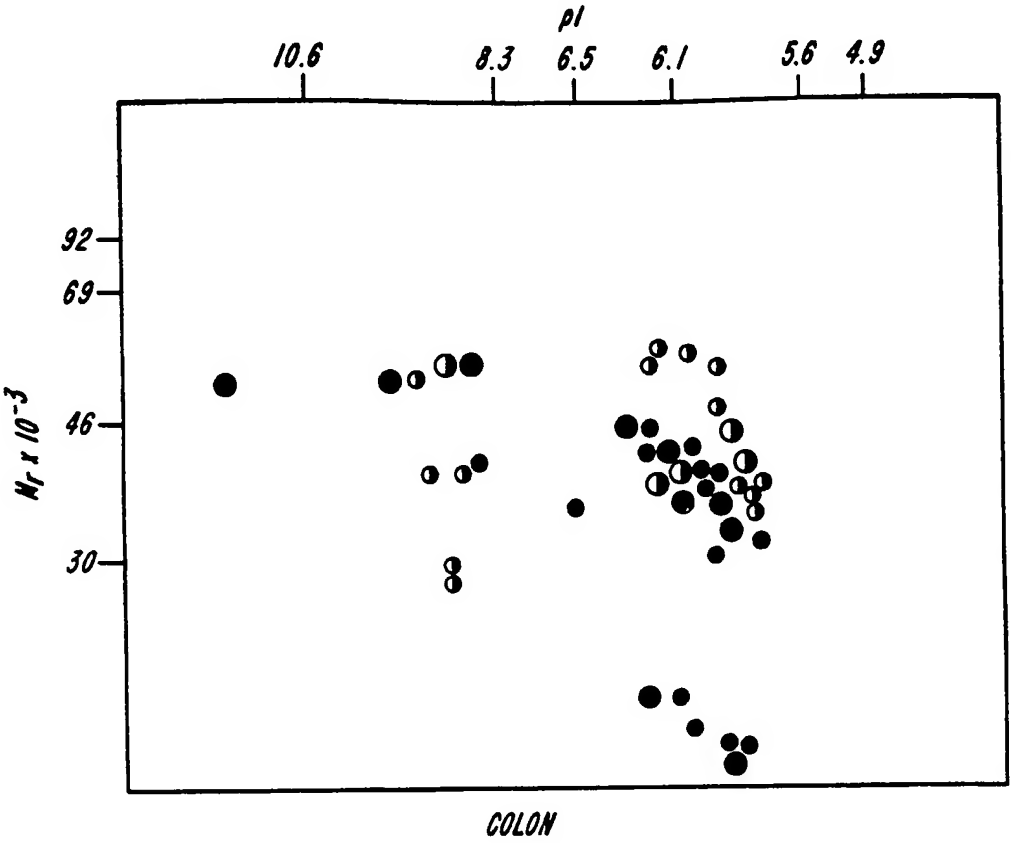


FIG. 2A



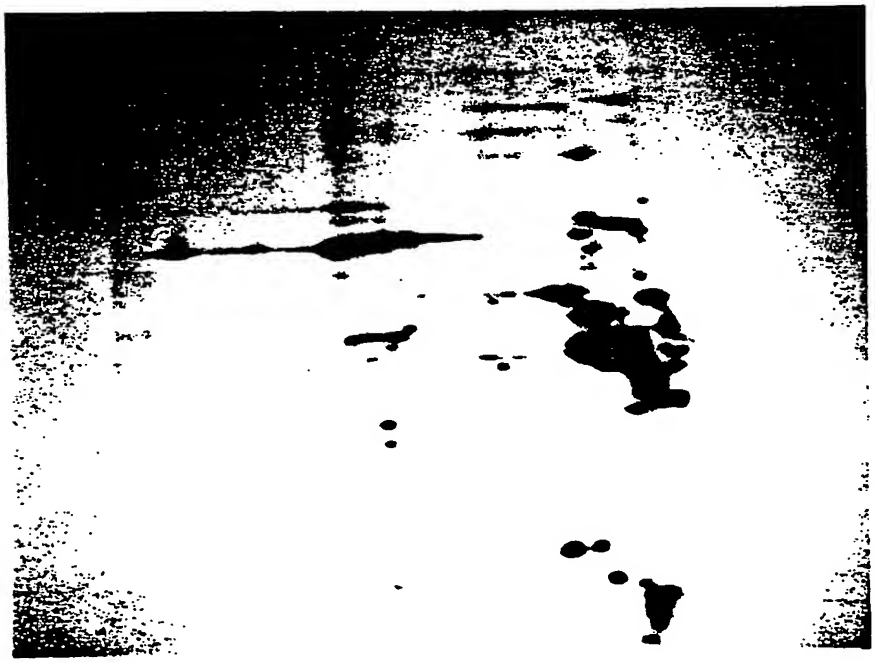
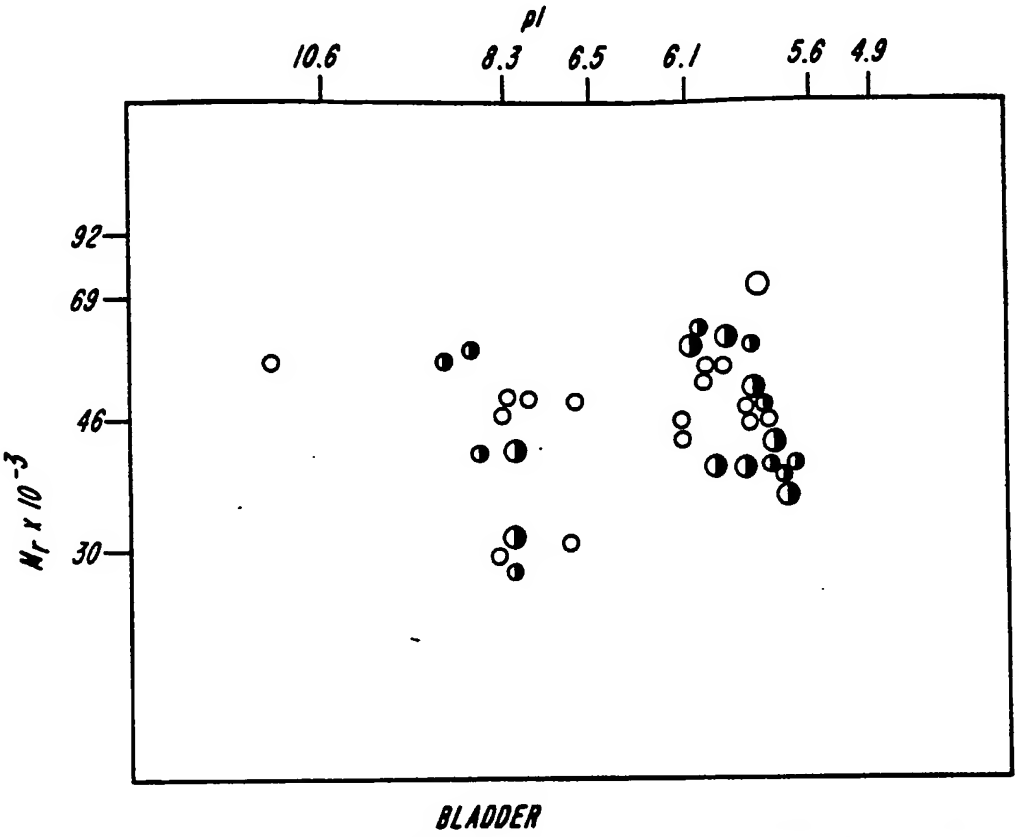


FIG. 2B



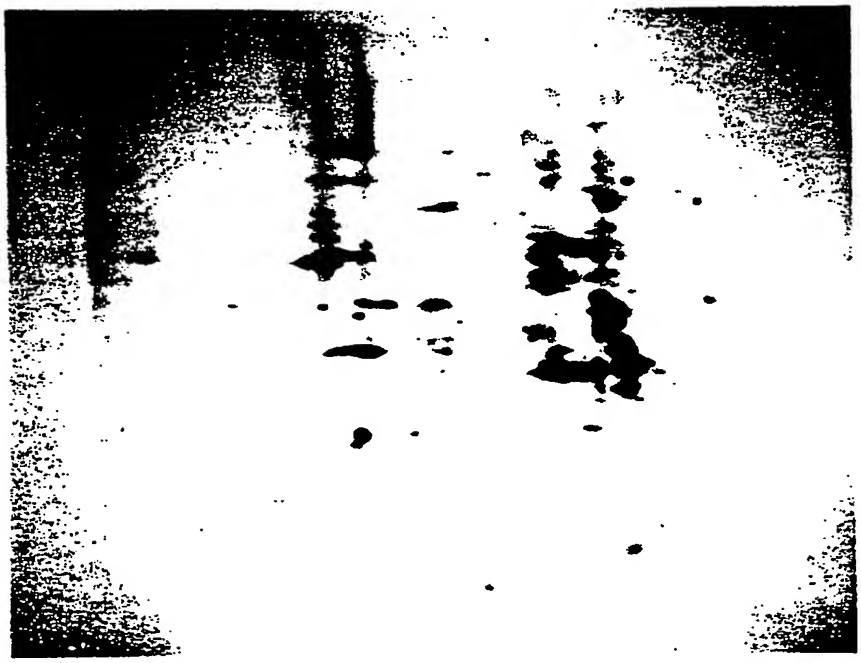


FIG. 2B

INTERNATIONAL SEARCH REPORT

International Application No PCT/US86/02798

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [3]

According to International Patent Classification (IPC) or to both National Classification and IPC

US Cl: 435/6,7,29;
Int. Cl. (4): C 12 Q 1/02, 1/68; See Attachment.

II. FIELDS SEARCHED

Minimum Documentation Searched [4]

Classification System	Classification Symbols	
U.S.	435/6,7,29,803 530/350,358,387,412,420 536/27	436/63,64,548,811,813 935/78

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [5]

COMPUTER SEARCH: CHEMICAL ABSTRACTS 1967-1987;
BIOSIS 1977-1987

III. DOCUMENTS CONSIDERED TO BE RELEVANT [14]

Category *	Citation of Document, [16] with indication, where appropriate, of the relevant passages [17]	Relevant to Claim No. [18]
X Y	Journal of Cellular Biochemistry, Volume 24, 1-14, issued 1984 (New York, New York, USA), R.M. FRANKLIN, "A Monoclonal Antibody Recognizes An Epitope Common To An Avian-Specific Nuclear Antigen And To Cytokeratins", see page 4, lines 10-25, page 11, last 6 lines and page 12, lines 1-6	17,18 1-16, 19-22
X Y	Chemical Abstracts, Volume 99, No. 23, issued 05 December 1983 (Columbus, Ohio, USA), W.J. HABETS, "Antibodies Against Distinct Nuclear Matrix Proteins Are Characteristic For Mixed connective Tissue Disease", see page 593-594, column 2, the abstract No. 192945a, Clin. Exp. Immunol. 1983, 54 (1), 265-76, (Eng).	17,18 1-16, 19-22

* Special categories of cited documents: [15]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the International filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search [2]	Date of Mailing of this International Search Report [2]
10 March 1987	20 MAR 1987
International Searching Authority [1]	**Signature of Authorized Officer [20]**
ISA/US	Esther M. Kepplinger

Form PCT/ISA/210 (second sheet) (May 1986)

PCT/US86/02798

I. CLASSIFICATION OF SUBJECT MATTER CONTINUED.

US Cl: 530/358, 387; 536/27
Int. Cl. (4): G 01 N 33/53; A 61 K 37/10, 39/395;
C 07 H 15/12

International Application No. PCT/US86/02798

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, [16] with indication, where appropriate, of the relevant passages [17]	Relevant to Claim No [18]
X Y	Chemical Abstracts, Volume 101, No. 3, issued 16 July 1984, (Columbus, Ohio, USA), I.B. ZBARSKII, "Nuclear Skeleton Substructure In Some Normal And Malignant Tissues", see page 245, column 1, the abstract No. 19258r, Macromol. Funct. Cell, Sov. - Ital. Symp., 2nd 1980 (Pub. 1982), 1, 114-23 (Eng).	17,18 1-16, 19-22
Y	US, A, 4,358,535 (FALKOW) 09 November 1982, see column 2, lines 43-51, column 3, lines 11-16 and column 6, lines 49-66.	7-9, 11-13, 15 19-22
Y,P	US, A, 4,569,916 (PENMAN) 11 February 1986, see column 5, lines 42-58.	1-22

International Application No. PCT/US86/02798

FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

X Y	Chemical Abstracts, Volume 101, No. 25, issued 17 December 1984 (Columbus, Ohio, USA), W.N. SCHMIDT, "Nuclear Matrix Antigens In Azo Dye-Induced Primary Rat Hepatomas", see page 549, column 2, the abstract No. 227908e, Cancer Res. 1984, 44(11), 5291-304 (Eng).	17,18 1-16, 19-22

V. ☐ OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE [10]

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1. ☐ Claim numbers .. .. , because they relate to subject matter [12] not required to be searched by this Authority, namely:

2. ☐ Claim numbers . ... , because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out [13], specifically:

VI. ☐ OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING [11]

This International Searching Authority found multiple inventions in this international application as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.

2. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4. ☐ As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest

☐ The additional search fees were accompanied by applicant's protest.

☐ No protest accompanied the payment of additional search fees.